



(12)发明专利申请

(10)申请公布号 CN 111429038 A

(43)申请公布日 2020.07.17

(21)申请号 202010336141.6

(22)申请日 2020.04.25

(71)申请人 华南理工大学

地址 510640 广东省广州市天河区五山路
381号

(72)发明人 李捷 余涛

(74)专利代理机构 广州粤高专利商标代理有限
公司 44102

代理人 何淑珍 江裕强

(51) Int. Cl.

G06Q 10/06(2012.01)

G06Q 50/06(2012.01)

H02J 3/28(2006.01)

H02J 3/38(2006.01)

权利要求书7页 说明书15页 附图2页

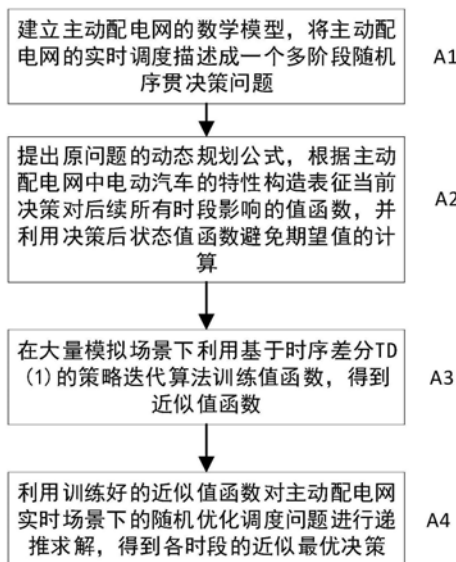
(54)发明名称

一种基于强化学习的主动配电网实时随机
优化调度方法

(57)摘要

本发明提出一种基于强化学习的主动配电网实时随机优化调度方法。所述方法包括以下步骤:建立主动配电网的数学模型;提出原问题的动态规划公式,根据主动配电网中电动汽车的特性构造表征当前决策对后续所有时段影响的值函数,并利用决策后状态值函数避免期望值的计算;训练值函数,得到近似值函数;利用训练好的近似值函数对主动配电网实时场景下的随机优化调度问题进行递推求解,得到各时段的近似最优决策。本发明在实时运行中时,一方面能更有效地提升调度收益,另一方面也起到“削峰填谷”的作用。当环境中的随机因素变化更加剧烈时,本发明依旧适用,能接近最优解,且实时调度过程中的计算时间并不受影响,具有较强的鲁棒性。

CN 111429038 A



1. 一种基于强化学习的主动配电网实时随机优化调度方法,其特征在于,包括以下步骤:

S1、建立主动配电网的数学模型,将主动配电网的实时调度描述成一个多阶段随机序贯决策问题;

S2、提出原问题的动态规划公式,根据主动配电网中电动汽车的特性构造表征当前决策对后续所有时段影响的值函数,并利用决策后状态值函数避免期望值的计算;

S3、在大量模拟场景下利用基于时序差分TD(1)的策略迭代算法训练值函数,得到近似值函数;

S4、利用训练好的近似值函数对主动配电网实时场景下的随机优化调度问题进行递推求解,得到各时段的近似最优决策。

2. 根据权利要求1所述的一种基于强化学习的主动配电网实时随机优化调度方法,其特征在于,步骤A1中,所述主动配电网的数学模型包括潮流约束、储能约束、电动汽车充电约束、分布式电源约束以及主动配电网实时调度的目标函数。

3. 根据权利要求2所述的一种基于强化学习的主动配电网实时随机优化调度方法,其特征在于,所述潮流约束如下:

$$\left\{ \begin{array}{l} P_{ij}(t) = k_{ij-1} \cdot \frac{(V_i(t) - V_j(t))}{x_{ij}} + k_{ij-2} \cdot \frac{(\delta_i(t) - \delta_j(t))}{x_{ij}} \\ Q_{ij}(t) = -k_{ij-1} \cdot \frac{(\delta_i(t) - \delta_j(t))}{x_{ij}} + k_{ij-2} \cdot \frac{(V_i(t) - V_j(t))}{x_{ij}} \\ P_i(t) = \sum_{j=1, j \neq i}^{NB} \frac{k_{ij-2}}{x_{ij}} (\delta_i(t) - \delta_j(t)) + \frac{k_{ij-1}}{x_{ij}} (V_i(t) - V_j(t)) ; \\ Q_i(t) = \sum_{j=1, j \neq i}^{NB} -\frac{k_{ij-1}}{x_{ij}} (\delta_i(t) - \delta_j(t)) + \frac{k_{ij-2}}{x_{ij}} (V_i(t) - V_j(t)) \\ k_{ij-1} = \frac{r_{ij} x_{ij}}{r_{ij}^2 + x_{ij}^2}, k_{ij-2} = \frac{x_{ij}^2}{r_{ij}^2 + x_{ij}^2} \\ -P_{ij, \max} \leq P_{ij} \leq P_{ij, \max}, V_{i, \min} \leq V_i \leq V_{i, \max} \end{array} \right. \quad (1)$$

式中, i, j 均为为节点编号, P_{ij}, Q_{ij}, r_{ij} 和 x_{ij} 分别为节点 i, j 间的有功传输功率、无功传输功率、电阻和电抗, P_i, Q_i, V_i 和 δ_i 分别为节点 i 的有功功率、无功功率、电压和相角, NB 为支路数; $P_{ij, \max}$ 和 $-P_{ij, \max}$ 分别为节点 i, j 间的最大和最小有功传输功率; $V_{i, \max}$ 和 $V_{i, \min}$ 分别为节点 i 的最大和最小电压;

除了与外部电网相连的节点,配电网内其它节点在 t 时段的功率如下:

$$P_{i,t} = P_{i,t}^{DG} + P_{i,t}^{ESS} - P_{i,t}^{CON} - \sum_{n=1}^{N_{EV}} P_{i,n,t}^{EV} ; \quad (2)$$

$$Q_{i,t} = Q_{i,t}^{DG} + Q_{i,t}^{ESS} - Q_{i,t}^{CON} - \sum_{n=1}^{N_{EV}} Q_{i,n,t}^{EV} ; \quad (3)$$

式中, $P_{i,t}^{DG}$ 和 $Q_{i,t}^{DG}$ 分别为 t 时段节点 i 上的分布式电源的有功和无功出力; $P_{i,t}^{ESS}$ 和 $Q_{i,t}^{ESS}$ 分别为 t 时段节点 i 上的储能的有功和无功出力; $P_{i,t}^{CON}$ 和 $Q_{i,t}^{CON}$ 为 t 时段节点 i 上的常规有功负

荷和无功负荷； $P_{i,n,t}^{EV}$ 和 $Q_{i,n,t}^{EV}$ 为t时段节点i上的电动汽车有功负荷和无功负荷；n为电动汽车编号， N_{EV} 为电动汽车数目；

所述储能约束如下：

$$E_{a,t+1} = \begin{cases} E_{a,t} - \frac{1}{\eta_a} P_{a,t} & P_{a,t} \geq 0 \\ E_{a,t} - \eta_a P_{a,t} & P_{a,t} < 0 \end{cases} ; \quad (4)$$

$$P_{a,t,\min} \leq P_{a,t} \leq P_{a,t,\max}; \quad (5)$$

$$E_{a,t,\min} \leq E_{a,t} \leq E_{a,t,\max}; \quad (6)$$

式(4)中，a为储能编号， $E_{a,t}$ 为t时段初第a个储能的能量， $E_{a,t+1}$ 为t时段末、t+1时段初第a个储能的能量； $P_{a,t}$ 为t时段第a个储能的充/放电功率，大于0表示放电，小于0表示充电；式(5)和(6)分别为功率和能量的上下限约束， $P_{a,t,\max}$ 和 $P_{a,t,\min}$ 分别为功率上下限， $E_{a,t,\max}$ 和 $E_{a,t,\min}$ 分别为能量上下限；

所述分布式电源约束如下：

$$\begin{cases} P_{d,t+1}^{DG} - P_{d,t}^{DG} \leq v_{up,d} \\ P_{d,t}^{DG} - P_{d,t+1}^{DG} \leq v_{down,d} \\ P_d^{DG,\min} \leq P_{d,t}^{DG} \leq P_d^{DG,\max} \end{cases} ; \quad (7)$$

其中，d为分布式电源编号， $P_{d,t+1}^{DG}$ 为第d台分布式电源在t+1时段中发出的有功功率； $v_{up,d}$ 和 $v_{down,d}$ 分别为第d台分布式电源的向上爬坡速率限制和向下爬坡速率限制， $P_d^{DG,\max}$ 和 $P_d^{DG,\min}$ 分别为第d台分布式电源输出功率的上限和下限；

所述单台电动汽车(EV)的充电约束如下：

$$\begin{cases} e_{e,t} = e_{e,t_{start}} + \eta \sum_{k=t_{start}}^t p_{e,k} \Delta t = e_{e,t-1} + \eta p_{e,t-1} \Delta t \\ e_{e,t_{end},\min} = e_{e,t_{end},\max} = e_{exp} \\ e_{e,t,\min} \leq e_{e,t} \leq e_{e,t,\max} \\ p_{e,t,\max} = \min(p_{max}, (e_{e,t,\max} - e_{e,t-1}) / \eta / \Delta t) \\ p_{e,t,\min} = \max(0, (e_{e,t,\min} - e_{e,t-1}) / \eta / \Delta t) \\ p_{e,t,\min} \leq p_{e,t} \leq p_{e,t,\max} \end{cases} ; \quad (8)$$

式中，e为电动汽车编号， $e_{e,t}$ 为t时段初第e台电动汽车电池的电量， η 为充电效率， $p_{e,k}$ 表示k时段第e台电动汽车的充电功率， $e_{e,t,\max}$ 、 $e_{e,t,\min}$ 分别为t时刻电动汽车能量的上下边界；能量的上边界表示电动汽车接入电网后便以最大功率充电充至 e_{exp} 这段时间内的电量变化曲线；能量的下边界表示接入后延迟充电直至离开的时刻恰好达到用户期望的电量值的电量变化曲线； e_{exp} 为用户希望离开时电池的电量，一般为电池的最大电量； $p_{e,t,\max}$ 、 $p_{e,t,\min}$ 分别为t时段第e台电动汽车充电功率的上下限约束， p_{max} 表示受充电桩和电动汽车本身性质影响的最大充电功率； Δt 为每个时段的时长； t_{end} 表示结束时段；

当电动汽车的渗透率过高时，将每台电动汽车作为决策变量进行求解会产生“维数灾”

的问题,可利用集群的思想,将一天分为T个时段,具有相同离开时段的电动汽车分为同一集群,将集群内单台电动汽车的充电模型叠加,可得到等效的集群模型:

$$\begin{cases} E_{\min,ec,t} \leq E_{ec,t} \leq E_{\max,ec,t} \\ E_{ec,t} = E_{ec,t_{\text{start}}} + \eta \sum_{k=t_{\text{start}}}^t P_{ec,k} \Delta t = E_{ec,t-1} + \eta P_{ec,t-1} \Delta t \\ P_{ec,t} = \sum_{e=1}^m P_{ec,e,t} \\ P_{\min,ec,t} = \sum_{e=1}^m P_{\min,ec,e,t} \leq P_{ec,t} \leq \sum_{e=1}^m P_{\max,ec,e,t} = P_{\max,ec,t} \\ E_{\min,ec,t} = \sum_{e=1}^m e_{\min,ec,e,t} \\ E_{\max,ec,t} = \sum_{e=1}^m e_{\max,ec,e,t} \end{cases}; \quad (9)$$

式中,ec为电动汽车集群编号, $E_{ec,t}$ 、 $E_{\max,ec,t}$ 、 $E_{\min,ec,t}$ 和 $P_{ec,t}$ 分别为t时段子集群ec的能量、能量上下界和子集群充电总功率; $P_{ec,e,t}$ 为子集群ec内第e辆电动汽车t时段的充电功率;m为t时段子集群ec的电动汽车数量;

所述目标函数,为考虑各类资源包括电动汽车、分布式电源的随机性下,通过实时调度决策,使得调度周期内的总收益最大,即将主动配电网的实时调度描述成一个多阶段随机序贯决策问题:

$$\max_{x_t \in \mathcal{X}_t} \mathbf{E} \left[\sum_{t=1}^T F_t \right]; \quad (10)$$

其中, x_t 表示决策变量, \mathcal{X}_t 表示可行域, $\mathbf{E}(\cdot)$ 表示取期望值运算, F_t 表示t时段的调度收益,为调度成本的相反数,具体如下:

$$F_t = - \left(C_t^{\text{trans}} + \sum_{d \in N_{\text{DG}}} C_{t,d}^{N_{\text{DG}}} + \sum_{a \in N_{\text{ESS}}} C_{t,a}^{N_{\text{ESS}}} \right); \quad (11)$$

其中, C_t^{trans} 、 $C_{t,a}^{N_{\text{ESS}}}$ 分别是t时段主动配电网向外部电网购电的成本、t时段第a个储能的运行成本, $C_{t,d}^{N_{\text{DG}}}$ 是t时段第d个分布式电源的运行成本,为二次函数; N_{DG} 、 N_{ESS} 分别表示分布式电源、储能的个数。

4. 根据权利要求1所述的一种基于强化学习的主动配电网实时随机优化调度方法,其特征在于,步骤S2中提出原问题的动态规划公式,即根据贝尔曼的最优性原理,可将多时段的优化决策问题转化成递推问题,将式(10)转化成:

$$V_t(S_t) = \max_{x_t \in \mathcal{X}} (F(S_t, x_t) + \xi E(V_{t+1}(S_{t+1} | S_t, x_t, W_t))); \quad (12)$$

其中, S_t 为主动配电网系统t时段所处的状态, $F(S_t, x_t)$ 为主动配电网系统在 S_t 状态时作了决策 x_t 后产生的实时收益, $V_t(S_t)$ 为主动配电网系统在 S_t 状态的值函数, $V_{t+1}(S_{t+1} | S_t, x_t, W_t)$ 为主动配电网系统在 S_t 状态的前提下,做出决策 x_t 、接收外源信息 W_t 后t+1时刻的值函数,值函数的意义为后续状态对当前状态的影响;

所述主动配电网系统t时段所处的状态 S_t 为

$$S_t = (R_{ec,t}^{\text{EV}}, R_{d,t}^{\text{DG}}, R_{a,t}^{\text{ESS}}, P_{i,t}, Q_{i,t}); \quad (13)$$

其中, $R_{ec,t}^{\text{EV}}$ 为t时段电动汽车子集群ec的可调资源量, $R_{ec,t}^{\text{EV}} = E_{\max,ec,t} - E_{ec,t}$,即t

时段电动汽车子集群 ec 的可调资源量为 t 时段电动汽车子集群 ec 的能量上界与该集群 ec 此时的能量之差,电动汽车子集群 ec 在 t 时段的能量为电池电量总和; $R_{d,t}^{DG}$ 为 t 时段分布式电源的可调资源量, $R_{d,t}^{DG} = P_d^{DG,max} - P_{d,t}^{DG}$,即 t 时段分布式电源的可调资源量为 t 时段分布式电源的最大发电功率与实际发电功率之差; $R_{ESSa,t}$ 为储能的可调容量,即剩余电量; $P_{i,t}$ 、 $Q_{i,t}$ 分别为 t 时段节点 i 的有功功率、无功功率;

决策变量 x_t 为:

$$x_t = (P_{ec,t}, P_{d,t}^{DG}, P_{a,t}); \quad (14)$$

外源信息 W_t 为:

$$W_t = (\tilde{w}_t^{EV}, \tilde{w}_t^{PV}, \tilde{w}_t^{WT}, \tilde{a}_t); \quad (15)$$

外源信息 W_t 表示 t 时段主动配电网系统更新的各随机变量的预测误差,其中 \tilde{w}_t^{EV} 表示因 t 时段电动汽车随机接入产生的误差,引起的变化包括各个电动汽车子集群内电动汽车的数量、集群能量轨迹上下边界、集群的充电功率极限; \tilde{w}_t^{PV} 、 \tilde{w}_t^{WT} 和 \tilde{a}_t 分别表示 t 时段光伏出力的误差、风机出力的误差和电价的误差。

5. 根据权利要求1所述的一种基于强化学习的主动配电网实时随机优化调度方法,其特征在于,步骤S2中根据主动配电网中电动汽车的特性构造表征当前决策对后续所有时段影响的值函数,并利用决策后状态值函数避免期望值的计算;

首先定义状态转移方程:

$$S_t = f(S_{t-1}, x_{t-1}, W_t); \quad (16)$$

表示 t 时段的主动配电网系统状态 S_t 由 $t-1$ 的状态 S_{t-1} 、 $t-1$ 时段所做的决策 x_{t-1} 、 t 时段的外源信息 W_t 共同决定;

接着将主动配电网系统状态 S_t 分为决策前状态 $S_t^{X^-}$ 和决策后状态 S_t^X 两个阶段;决策前,主动配电网系统接收外源信息 W_t ,改变状态, $S_t^{X^-} = S_{t-1}^X + f^{X^-}(W_t)$;做出决策后、未接受外源信息前,主动配电网系统的状态进一步变化 $S_t^X = S_t^{X^-} + f^X(x_t)$;

根据电动汽车特性进行设计,决策前每个电动汽车子集群的电池电量 $E_{ec,t}$ 不变,但观测到新的电动汽车的随机接入,即 \tilde{w}_t^{EV} ,导致电动汽车集群的能量上下边界 $E_{ec,t,max}$ 、 $E_{ec,t,min}$ 发生变化,从而电动汽车子集群的可调资源量由 $R_{ec,t-1}^X$ 变化至 $R_{ec,t}^X$;决策后,每个电动汽车集群的电池电量 $E_{ec,t}$ 根据所作的决策发生改变,可调资源量变成 $R_{Xec,t}$,实现主动配电网系统状态的转移;

根据前述内容将主动配电网系统状态分成决策前与决策后状态,则式(12)可分成两部分,得到决策前状态值函数 $V_t^{X^-}(S_t^{X^-}) = \max_{x_t \in X} (F(S_t, x_t) + \xi V_t^X(S_t^X | S_t^{X^-}))$ 和决策后状态值函数 $V_t^X(S_t^X) = E(V_{t+1}^{X^-}(S_{t+1}^{X^-} | S_t^X))$;注意到决策前状态值函数是确定性问题,因此只要解决决策后状态值函数 $V_t^X(S_t^X)$ 的近似,即可通过解决确定性问题来获得最优解。

6. 根据权利要求1所述的一种基于强化学习的主动配电网实时随机优化调度方法,其特征在于,步骤S3中,在大量模拟场景下利用基于时序差分TD(1)的策略迭代算法训练值函数,得到收敛的近似值函数,步骤如下:

S3.1、初始化 $v_t^n(r, W_t)$, $v_t^n(r, W_t)$ 为第n次迭代中t时段第r段的斜率,直接令 $v_t^0(r, W_1) = v_t^0(r, W_2) = \dots = v_t^0(r, W_t) = \dots = v_t^0(r, W_T) = 0$; 生成n个训练样本 TS^1, TS^2, \dots, TS^n ; 令迭代次数 $g=1$;

S3.2、从 $t=1, \dots, T$, 进行前推计算;

S3.3、从 $t=T, \dots, 1$, 进行回代计算;

S3.4、令 $g=g+1$, 返回步骤S2; 当 $g>n$ 时迭代停止, 得到收敛的各分段线性凹函数的斜率。

7. 根据权利要求6所述的一种基于强化学习的主动配电网实时随机优化调度方法,其特征在于,步骤S3.2的具体步骤如下:

S3.2.1、令 $t=1$;

S3.2.2、根据最新的随机变量更新主动配电网系统状态并利用上一次迭代后的各分段斜率 $v_t^{n-1}(r, W_t)$ 求得各决策变量 x_t^n , 决策后的主动配电网系统状态 $S_t^{n,X}$ 包括决策后的可调容量 $R_t^{n,X}$;

S3.2.3、若 $t>1$, 对每个可调容量施加正负扰动, 并计算其正负边际效益, 第n次迭代中的第i个可调容量 $R_{ec,t}^{n,X}$ 的正负边际效益如下:

$$\begin{cases} \hat{F}_{ec,t}^{n+} = \frac{1}{\delta^+} (F_t(S_t^{n+}, x_t^{n+}) - F_t(S_t, x_t)) \\ \hat{F}_{ec,t}^{n-} = \frac{1}{\delta^-} (F_t(S_t, x_t) - F_t(S_t^{n-}, x_t^{n-})) \end{cases}; \quad (19)$$

式中, $\hat{F}_{ec,t}^{n+}$ 、 $\hat{F}_{ec,t}^{n-}$ 分别为第ec个电动汽车子集群可调容量的正负边际效益, δ^+ 、 δ^- 分别为正负扰动量, F_t 为立即成本; S_t^{n+} 、 S_t^{n-} 分别为t时段施加正、负扰动后的主动配电网系统状态, x_t^{n+} 、 x_t^{n-} 分别为t时段施加正、负扰动后所做的决策;

S3.2.4、计算每个可调容量施加正负扰动后的资源变化量:

$$\begin{cases} \Delta R_{ec,t}^{n+,X} = \frac{1}{\delta^+} (R_{ec,t}^{n+,X} - R_{ec,t}^{n,X}) \\ \Delta R_{ec,t}^{n-,X} = \frac{1}{\delta^-} (R_{ec,t}^{n,X} - R_{ec,t}^{n-,X}) \end{cases}; \quad (20)$$

其中, $\Delta R_{ec,t}^{n+,X}$ 为施加正扰动后的资源变化量, $\Delta R_{ec,t}^{n-,X}$ 为施加负扰动后的资源变化量; $R_{ec,t}^{n+,X}$ 、 $R_{ec,t}^{n-,X}$ 分别为施加正、负扰动后的资源量;

S3.2.5、令 $t=t+1$, 若 $t \leq T$, 返回步骤S2.2; 若 $t > T$, 循环终止。

8. 根据权利要求6所述的一种基于强化学习的主动配电网实时随机优化调度方法,其特征在于,步骤S3.3的具体步骤如下:

S3.3.1、令 $t=T$;

S3.3.2、计算本次迭代过程中决策前各分段斜率的正负边际值,正边际值的计算:

$$\hat{v}_t^{n+}(r, S_t^n) = \begin{cases} \hat{F}_{ec,t}^{n+} + R_{ec,t}^{n+,X} \hat{v}_{t+1}^{n+}(r, S_{t+1}^n), & t \neq T \\ \hat{F}_{ec,T}^{n+}, & t = T \end{cases}; \quad (21)$$

负边际值的计算:

$$\hat{v}_t^{n-}(r, S_t^n) = \begin{cases} \hat{F}_{ec,t}^{n-} + R_{ec,t}^{n-,X} \hat{v}_{t+1}^{n-}(r, S_{t+1}^n), & t \neq T \\ \hat{F}_{ec,T}^{n-}, & t = T \end{cases}; \quad (22)$$

根据 $\hat{v}_t^{n+}(r, S_t^n)$ 、 $\hat{v}_t^{n-}(r, S_t^n)$ 以及上一次迭代的决策后边际值 $\hat{v}_t^{(n-1)+}(r, S_t^{n-1,X})$ 、 $\hat{v}_t^{(n-1)-}(r, S_t^{n-1,X})$ 求得本次迭代的决策后边际值:

$$\begin{cases} \hat{v}_{t-1}^{n+}(r, S_{t-1}^{n,X}) = (1-\alpha)\hat{v}_{t-1}^{(n-1)+}(r, S_{t-1}^{n-1,X}) + \alpha\hat{v}_t^{n+}(r, S_t^n) \\ \hat{v}_{t-1}^{n-}(r, S_{t-1}^{n,X}) = (1-\alpha)\hat{v}_{t-1}^{(n-1)-}(r, S_{t-1}^{n-1,X}) + \alpha\hat{v}_t^{n-}(r, S_t^n) \end{cases}; \quad (23)$$

α 为迭代步长,取值范围为0~1;

S3.3.3、根据 $\hat{v}_{t-1}^{n+}(r, S_{t-1}^{n,X})$ 、 $\hat{v}_{t-1}^{n-}(r, S_{t-1}^{n,X})$,采用凹自适应值估计法(TheConcaveAdaptiveValueEstimationAlgorithm,CAVE)对各分段线性凹函数的斜率进行更新;

S3.3.4、令 $t=t-1$,若 $t \geq 1$,返回步骤S3.2;若 $t < 1$,循环终止。

9.根据权利要求6所述的一种基于强化学习的主动配电网实时随机优化调度方法,其特征在于,步骤S3.4中,利用分段线性函数法将决策后状态值函数用分段线性凹函数表示,通过不断迭代更新得到收敛的分段线性凹函数的斜率,从而得到近似的决策后值函数即近似值函数;

所述分段线性函数法将决策后状态值函数用分段线性凹函数表示步骤如下:将可调资源用断点 $R = [\rho, 2\rho, \dots, \beta\rho]$ 表示, R 表示断点集合, β 代表段数, ρ 为每段的长度,将分段线性凹函数的斜率向量用有限集合 $v_t(W_t) = [v_t(1, W_t), v_t(2, W_t), \dots, v_t(\beta, W_t)]$ 表示, $v_t(W_t)$ 表示斜率集合, $v_t(\beta, W_t)$ 表示第 β 段的斜率, $[v_t(1, W_t) \geq v_t(2, W_t), \geq \dots \geq v_t(\beta, W_t)]$,保证斜率单调递减;则式(12)可进一步转化为:

$$x_t = \arg \max_{x_t \in X} (F(S_t, x_t) + \xi \sum_{r=1}^{\beta} v_t(r, W_t) y_{tr}); \quad (17)$$

表示当决策 x_t 使得上式取最大值时,即为最优决策;须满足:

$$\begin{cases} \sum_{r=1}^{\beta} y_{tr} = R_{ec,t}^X \\ 0 \leq y_{tr} \leq \rho, r \in \{1, \dots, \beta\} \end{cases}; \quad (18)$$

其中, r 表示第 r 段线性凹函数, ρ 为每段线性凹函数的长度, y_{tr} 为每段线性凹函数的资

源量,其中,假设所有的分段线性凹函数的长度都是均匀的。

10.根据权利要求1所述的一种基于强化学习的主动配电网实时随机优化调度方法,其特征在于,步骤S4中,利用训练好的近似值函数对主动配电网实时场景下的随机优化调度问题进行递推求解,得到各时段的近似最优决策,步骤如下:

S4.1、令 $t=1$;

S4.2、更新当前时段的外源信息,包括新接入的电动汽车、光伏出力的误差、风机出力的误差以及电价的误差;

S4.3、利用步骤S3中训练好的近似值函数,根据式(17)计算出 t 时段最优决策;

S4.4、令 $t=t+1$,若 $t \leq T$,返回步骤S4.2;若 $t > T$,循环终止,得到各时段的近似最优决策。

一种基于强化学习的主动配电网实时随机优化调度方法

技术领域

[0001] 本发明涉及电力系统主动配电网优化调度领域,尤其涉及一种基于强化学习的主动配电网实时随机优化调度方法。

背景技术

[0002] 随着分布式电源(Distributed Generator, DG)、储能装置(Energy Storage, ES)和柔性负荷(Flexible Load, FL)广泛接入电网,传统配电网正逐渐向主动配电网(Active Distribution Network, ADN)转变。不同于传统配电网的能量单向流动,ADN中的“主动”二字体现在能量的双向流动,即不止主网能向配网传输功率,配网中的DG、ES、FL也可以反过来向主网传输功率。ADN的不断发展给电力系统的安全经济运行带来一系列的挑战。已有大量文献对ADN的调度方法进行了研究,提供了多种方法促进配电网中分布式资源的消纳,但是并未考虑随机性的影响(顾伟,任佳依,高君,高菲,宋晓辉,刘海波.含分布式电源和可调负荷的售电公司优化调度模型[J].电力系统自动化,2017,41(14):37-44.肖浩,裴玮,孔力.含大规模电动汽车接入的主动配电网多目标优化调度方法[J].电工技术学报,2017,32(S2):179-189.Hu J, Cong H, Jiang C. Coordinated scheduling model of power system with active distribution networks based on multi-agent system[J]. Journal of Modern Power Systems and Clean Energy, 2018, 6(3): 521-531. 欧阳聪, 刘明波, 林舜江, 冯汉中. 采用同步型交替方向乘子法的微电网分散式动态经济调度算法[J]. 电工技术学报, 2017, 32(05): 134-142.) ; 针对考虑随机性的优化问题, 鲁棒优化(梁俊文, 林舜江, 刘明波, 宋雨浓, 范官盛, 何森, 蒋浩. 主动配电网分布式鲁棒优化调度方法[J]. 电网技术, 2019, 43(04): 1336-1344.)、机会约束(周星球, 郑凌蔚, 杨兰, 仇琦. 考虑多重不确定性的综合能源系统日前优化调度[J/OL]. 电网技术: 1-8[2020-04-13]. <https://doi.org/10.13335/j.1000-3673.pst.2019.2169>.)等算法得到了广泛应用。但大部分文献是站在日前调度决策的角度考虑随机性的影响, 仅仅依赖日前调度方案, 并不能实现经济效益的最大化, 当预测误差较大时甚至可能出现安全运行问题。

[0003] 强化学习可以解决考虑随机性影响下的多时段决策问题, 在电力系统优化运行领域得到广泛关注, 针对含储能或抽水蓄能站的电网动态经济调度问题, 提出存储器模型以处理随机性(SHUI Hang, FANG Jiakun, AI Xiaomeng, et al. Stochastic optimization of economic dispatch for microgrid based on approximate dynamic programming [J]. IEEE Transactions on Smart Grid, 2018 (Early Access) 简淦杨, 刘明波, 林舜江. 随机动态经济调度问题的存储器建模及近似动态规划算法[J]. 中国电机工程学报, 2014, 34(25): 4333-4340.)。但上述方法针对单个储能设备进行值函数的设计, 并无法简单应用到数量巨大的电动汽车中去。在计算效率方面, 利用逐次投影近似法更新值函数(刘翠平, 林舜江, 刘明波, 简淦杨, 陆文甜. 应用近似动态规划算法求解安全约束随机动态经济调度问题[J]. 电力系统自动化, 2016, 40(22): 34-42. SALAS D F, POWELL W B. Benchmarking a scalable approximate dynamic programming algorithm for stochastic control of

grid-level energy storage[J].Informs Journal on Computing,2018,30(1):106-123.),并采用投影运算的方法保证斜率单调递减,从而恢复值函数的凹性,但该方法收敛速度较慢,且对于初始斜率的依赖性过高。

[0004] 因此,迫切需要找到一种适应电动汽车等新能源出力特点下的主动配电网实时随机优化调度方法。

发明内容

[0005] 本发明提出一种基于强化学习的主动配电网实时随机优化调度方法。建立主动配电网的数学模型,将主动配电网的实时调度描述成一个多阶段随机序贯决策问题;提出原问题的动态规划公式,根据主动配电网中电动汽车的特性构造表征当前决策对后续所有时段影响的值函数,并利用决策后状态值函数避免期望值的计算;在大量模拟场景下利用基于时序差分TD(1)的策略迭代算法训练值函数,得到近似值函数;利用训练好的近似值函数对主动配电网实时场景下的随机优化调度问题进行递推求解,得到各时段的近似最优决策。

[0006] 本发明的目的至少通过如下技术方案之一实现。

[0007] 一种基于强化学习的主动配电网实时随机优化调度方法,包括以下步骤:

[0008] S1、建立主动配电网的数学模型,将主动配电网的实时调度描述成一个多阶段随机序贯决策问题;

[0009] S2、提出原问题的动态规划公式,根据主动配电网中电动汽车的特性构造表征当前决策对后续所有时段影响的值函数,并利用决策后状态值函数避免期望值的计算;

[0010] S3、在大量模拟场景下利用基于时序差分TD(1)的策略迭代算法训练值函数,得到近似值函数;

[0011] S4、利用训练好的近似值函数对主动配电网实时场景下的随机优化调度问题进行递推求解,得到各时段的近似最优决策。

[0012] 进一步地,步骤A1中,所述主动配电网的数学模型包括潮流约束、储能约束、电动汽车充电约束、分布式电源约束以及主动配电网实时调度的目标函数。

[0013] 进一步地,所述潮流约束如下:

$$\begin{cases}
 P_{ij}(t) = k_{ij-1} \cdot \frac{(V_i(t) - V_j(t))}{x_{ij}} + k_{ij-2} \cdot \frac{(\delta_i(t) - \delta_j(t))}{x_{ij}} \\
 Q_{ij}(t) = -k_{ij-1} \cdot \frac{(\delta_i(t) - \delta_j(t))}{x_{ij}} + k_{ij-2} \cdot \frac{(V_i(t) - V_j(t))}{x_{ij}} \\
 P_i(t) = \sum_{j=1, j \neq i}^{NB} \frac{k_{ij-2}}{x_{ij}} (\delta_i(t) - \delta_j(t)) + \frac{k_{ij-1}}{x_{ij}} (V_i(t) - V_j(t)) ; \\
 Q_i(t) = \sum_{j=1, j \neq i}^{NB} -\frac{k_{ij-1}}{x_{ij}} (\delta_i(t) - \delta_j(t)) + \frac{k_{ij-2}}{x_{ij}} (V_i(t) - V_j(t)) \\
 k_{ij-1} = \frac{r_{ij} x_{ij}}{r_{ij}^2 + x_{ij}^2}, k_{ij-2} = \frac{x_{ij}^2}{r_{ij}^2 + x_{ij}^2} \\
 -P_{ij, \max} \leq P_{ij} \leq P_{ij, \max}, V_{i, \min} \leq V_i \leq V_{i, \max}
 \end{cases} \quad (1)$$

[0015] 式中, i, j 均为节点编号, P_{ij}, Q_{ij}, r_{ij} 和 x_{ij} 分别为节点 i, j 间的有功传输功率、无功传输功率、电阻和电抗, P_i, Q_i, V_i 和 δ_i 分别为节点 i 的有功功率、无功功率、电压和相角, NB 为支路数; $P_{ij, \max}$ 和 $-P_{ij, \max}$ 分别为节点 i, j 间的最大和最小有功传输功率; $V_{i, \max}$ 和 $V_{i, \min}$ 分别为节点 i 的最大和最小电压;

[0016] 除了与外部电网相连的节点, 配电网内其它节点在 t 时段的功率如下:

$$[0017] \quad P_{i,t} = P_{i,t}^{\text{DG}} + P_{i,t}^{\text{ESS}} - P_{i,t}^{\text{CON}} - \sum_{n=1}^{N_{\text{EV}}} P_{i,n,t}^{\text{EV}}; \quad (2)$$

$$[0018] \quad Q_{i,t} = Q_{i,t}^{\text{DG}} + Q_{i,t}^{\text{ESS}} - Q_{i,t}^{\text{CON}} - \sum_{n=1}^{N_{\text{EV}}} Q_{i,n,t}^{\text{EV}}; \quad (3)$$

[0019] 式中, $P_{i,t}^{\text{DG}}$ 和 $Q_{i,t}^{\text{DG}}$ 分别为 t 时段节点 i 上的分布式电源的有功和无功出力; $P_{i,t}^{\text{ESS}}$ 和 $Q_{i,t}^{\text{ESS}}$ 分别为 t 时段节点 i 上的储能的有功和无功出力; $P_{i,t}^{\text{CON}}$ 和 $Q_{i,t}^{\text{CON}}$ 为 t 时段节点 i 上的常规有功负荷和无功负荷; $P_{i,n,t}^{\text{EV}}$ 和 $Q_{i,n,t}^{\text{EV}}$ 为 t 时段节点 i 上的电动汽车有功负荷和无功负荷; n 为电动汽车编号, N_{EV} 为电动汽车数目;

[0020] 所述储能约束如下:

$$[0021] \quad E_{a,t+1} = \begin{cases} E_{a,t} - \frac{1}{\eta_a} P_{a,t} & P_{a,t} \geq 0 \\ E_{a,t} - \eta_a P_{a,t} & P_{a,t} < 0 \end{cases}; \quad (4)$$

$$[0022] \quad P_{a,t,\min} \leq P_{a,t} \leq P_{a,t,\max}; \quad (5)$$

$$[0023] \quad E_{a,t,\min} \leq E_{a,t} \leq E_{a,t,\max}; \quad (6)$$

[0024] 式 (4) 中, a 为储能编号, $E_{a,t}$ 为 t 时段初第 a 个储能的能量, $E_{a,t+1}$ 为 t 时段末、 $t+1$ 时段初第 a 个储能的能量; $P_{a,t}$ 为 t 时段第 a 个储能的充/放电功率, 大于 0 表示放电, 小于 0 表示充电; 式 (5) 和 (6) 分别为功率和能量的上下限约束, $P_{a,t,\max}$ 和 $P_{a,t,\min}$ 分别为功率上下限, $E_{a,t,\max}$ 和 $E_{a,t,\min}$ 分别为能量上下限;

[0025] 所述分布式电源约束如下:

$$[0026] \quad \begin{cases} P_{d,t+1}^{\text{DG}} - P_{d,t}^{\text{DG}} \leq v_{\text{up},d} \\ P_{d,t}^{\text{DG}} - P_{d,t+1}^{\text{DG}} \leq v_{\text{down},d} \\ P_d^{\text{DG},\min} \leq P_{d,t}^{\text{DG}} \leq P_d^{\text{DG},\max} \end{cases}; \quad (7)$$

[0027] 其中, d 为分布式电源编号, $P_{d,t+1}^{\text{DG}}$ 为第 d 台分布式电源在 $t+1$ 时段中发出的有功功率; $v_{\text{up},d}$ 和 $v_{\text{down},d}$ 分别为第 d 台分布式电源的向上爬坡速率限制和向下爬坡速率限制, $P_d^{\text{DG},\max}$ 和 $P_d^{\text{DG},\min}$ 分别为第 d 台分布式电源输出功率的上限和下限;

[0028] 所述单台电动汽车 (EV) 的充电约束如下:

$$[0029] \quad \begin{cases} e_{e,t} = e_{e,t_{\text{start}}} + \eta \sum_{k=t_{\text{start}}}^t p_{e,k} \Delta t = e_{e,t-1} + \eta p_{e,t-1} \Delta t \\ e_{e,t_{\text{end}},\text{min}} = e_{e,t_{\text{end}},\text{max}} = e_{\text{exp}} \\ e_{e,t,\text{min}} \leq e_{e,t} \leq e_{e,t,\text{max}} \\ p_{e,t,\text{max}} = \min(p_{\text{max}}, (e_{e,t,\text{max}} - e_{e,t-1}) / \eta / \Delta t) \\ p_{e,t,\text{min}} = \max(0, (e_{e,t,\text{min}} - e_{e,t-1}) / \eta / \Delta t) \\ p_{e,t,\text{min}} \leq p_{e,t} \leq p_{e,t,\text{max}} \end{cases}; \quad (8)$$

[0030] 式中, e 为电动汽车编号, $e_{e,t}$ 为 t 时段初第 e 台电动汽车电池的电量, η 为充电效率, $p_{e,k}$ 表示 k 时段第 e 台电动汽车的充电功率, $e_{e,t,\text{max}}$ 、 $e_{e,t,\text{min}}$ 分别为 t 时刻电动汽车能量的上下边界; 能量的上边界表示电动汽车接入电网后便以最大功率充电充至 e_{exp} 这段时间内的电量变化曲线; 能量的下边界表示接入后延迟充电直至离开的时刻恰好达到用户期望的电量值的电量变化曲线; e_{exp} 为用户希望离开时电池的电量, 一般为电池的最大电量; $p_{e,t,\text{max}}$ 、 $p_{e,t,\text{min}}$ 分别为 t 时段第 e 台电动汽车充电功率的上下限约束, p_{max} 表示受充电桩和电动汽车本身性质影响的最大充电功率; Δt 为每个时段的时长; t_{end} 表示结束时段;

[0031] 当电动汽车的渗透率过高时, 将每台电动汽车作为决策变量进行求解会产生“维数灾”的问题, 可利用集群的思想, 将一天分为 T 个时段, 具有相同离开时段的电动汽车分为同一集群, 将集群内单台电动汽车的充电模型叠加, 可得到等效的集群模型:

$$[0032] \quad \begin{cases} E_{\text{min},ec,t} \leq E_{ec,t} \leq E_{\text{max},ec,t} \\ E_{ec,t} = E_{ec,t_{\text{start}}} + \eta \sum_{k=t_{\text{start}}}^t P_{ec,k} \Delta t = E_{ec,t-1} + \eta P_{ec,t-1} \Delta t \\ P_{ec,t} = \sum_{e=1}^m p_{ec,e,t} \\ P_{\text{min},ec,t} = \sum_{e=1}^m p_{\text{min},ec,e,t} \leq P_{ec,t} \leq \sum_{e=1}^m p_{\text{max},ec,e,t} = P_{\text{max},ec,t} \\ E_{\text{min},ec,t} = \sum_{e=1}^m e_{\text{min},ec,e,t} \\ E_{\text{max},ec,t} = \sum_{e=1}^m e_{\text{max},ec,e,t} \end{cases}; \quad (9)$$

[0033] 式中, ec 为电动汽车集群编号, $E_{ec,t}$ 、 $E_{\text{max},ec,t}$ 、 $E_{\text{min},ec,t}$ 和 $P_{ec,t}$ 分别为 t 时段子集群 ec 的能量、能量上下界和子集群充电总功率; $p_{ec,e,t}$ 为子集群 ec 内第 e 辆电动汽车 t 时段的充电功率; m 为 t 时段子集群 ec 的电动汽车数量;

[0034] 所述目标函数, 为考虑各类资源包括电动汽车、分布式电源的随机性下, 通过实时调度决策, 使得调度周期内的总收益最大, 即将主动配电网的实时调度描述成一个多阶段随机序贯决策问题:

$$[0035] \quad \max_{x_t \in X_t} \mathbf{E} \left[\sum_{t=1}^T F_t \right]; \quad (10)$$

[0036] 其中, x_t 表示决策变量, X_t 表示可行域, $\mathbf{E}(\cdot)$ 表示取期望值运算, F_t 表示 t 时段的调度收益, 为调度成本的相反数, 具体如下:

$$[0037] \quad F_t = - \left(C_t^{\text{trans}} + \sum_{d \in N_{\text{DG}}} C_{t,d}^{\text{NDG}} + \sum_{a \in N_{\text{ESS}}} C_{t,a}^{\text{NESS}} \right); \quad (11)$$

[0038] 其中, C_t^{trans} 、 $C_{t,a}^{\text{NESS}}$ 分别是 t 时段主动配电网向外部电网购电的成本、 t 时段第 a 个储

能的运行成本, $C_{t,d}^{NDG}$ 是t时段第d个分布式电源的运行成本, 为二次函数; N_{DG}, N_{ESS} 分别表示分布式电源、储能的个数。

[0039] 进一步地, 步骤S2中提出原问题的动态规划公式, 即根据贝尔曼的最优性原理, 可将多时段的优化决策问题转化成递推问题, 将式 (10) 转化成:

$$[0040] \quad V_t(S_t) = \max_{x_t \in X} (F(S_t, x_t) + \xi E(V_{t+1}(S_{t+1} | S_t, x_t, W_t))); \quad (12)$$

[0041] 其中, S_t 为主动配电网系统t时段所处的状态, $F(S_t, x_t)$ 为主动配电网系统在 S_t 状态时作了决策 x_t 后产生的实时收益, $V_t(S_t)$ 为主动配电网系统在 S_t 状态的值函数, $V_{t+1}(S_{t+1} | S_t, x_t, W_t)$ 为主动配电网系统在 S_t 状态的前提下, 做出决策 x_t 、接收外源信息 W_t 后t+1时刻的值函数, 值函数的意义为后续状态对当前状态的影响;

[0042] 所述主动配电网系统t时段所处的状态 S_t 为

$$[0043] \quad S_t = (R_{ec,t}^{EV}, R_{d,t}^{DG}, R_{a,t}^{ESS}, P_{i,t}, Q_{i,t}); \quad (13)$$

[0044] 其中, $R_{ec,t}^{EV}$ 为t时段电动汽车子集群ec的可调资源量, $R_{ec,t}^{EV} = E_{\max,ec,t} - E_{ec,t}$, 即t时段电动汽车子集群ec的可调资源量为t时段电动汽车子集群ec的能量上界与该集群ec此时的能量之差, 电动汽车子集群ec在t时段的能量为电池电量总和; $R_{d,t}^{DG}$ 为t时段分布式电源的可调资源量, $R_{d,t}^{DG} = P_d^{DG,\max} - P_{d,t}^{DG}$, 即t时段分布式电源的可调资源量为t时段分布式电源的最大发电功率与实际发电功率之差; $R_{ESSa,t}$ 为储能的可调容量, 即剩余电量; $P_{i,t}, Q_{i,t}$ 分别为t时段节点i的有功功率、无功功率;

[0045] 决策变量 x_t 为:

$$[0046] \quad x_t = (P_{ec,t}, P_{d,t}^{DG}, P_{a,t}); \quad (14)$$

[0047] 外源信息 W_t 为:

$$[0048] \quad W_t = (\tilde{w}_t^{EV}, \tilde{w}_t^{PV}, \tilde{w}_t^{WT}, \tilde{a}_t); \quad (15)$$

[0049] 外源信息 W_t 表示t时段主动配电网系统更新的各随机变量的预测误差, 其中 \tilde{w}_t^{EV} 表示因t时段电动汽车随机接入产生的误差, 引起的变化包括各个电动汽车子集群内电动汽车的数量、集群能量轨迹上下边界、集群的充电功率极限; \tilde{w}_t^{PV} 、 \tilde{w}_t^{WT} 和 \tilde{a}_t 分别表示t时段光伏出力的误差、风机出力的误差和电价的误差。

[0050] 进一步地, 步骤S2中根据主动配电网中电动汽车的特性构造表征当前决策对后续所有时段影响的值函数, 并利用决策后状态值函数避免期望值的计算;

[0051] 首先定义状态转移方程:

$$[0052] \quad S_t = f(S_{t-1}, x_{t-1}, W_t); \quad (16)$$

[0053] 表示t时段的主动配电网系统状态 S_t 由t-1的状态 S_{t-1} 、t-1时段所做的决策 x_{t-1} 、t时段的外源信息 W_t 共同决定;

[0054] 接着将主动配电网系统状态 S_t 分为决策前状态 S_t^X 和决策后状态 S_t^Y 两个阶段; 决策

前,主动配电网系统接收外源信息 W_t ,改变状态, $S_t^{X^-} = S_{t-1}^{X^-} + f^{X^-}(W_t)$;做出决策后、未接受外源信息前,主动配电网系统的状态进一步变化 $S_t^X = S_t^{X^-} + f^X(x_t)$;

[0055] 根据电动汽车特性进行设计,决策前每个电动汽车子集群的电池电量 $E_{ec,t}$ 不变,但观测到新的电动汽车的随机接入,即 \tilde{w}_t^{EV} ,导致电动汽车集群的能量上下边界 $E_{ec,t,max}$ 、 $E_{ec,t,min}$ 发生变化,从而电动汽车子集群的可调资源量由 $R_{ec,t-1}^X$ 变化至 $R_{ec,t}^X$;决策后,每个电动汽车集群的电池电量 $E_{ec,t}$ 根据所作的决策发生改变,可调资源量变成 $R_{Xec,t}$,实现主动配电网系统状态的转移;

[0056] 根据前述内容将主动配电网系统状态分成决策前与决策后状态,则式(12)可分成两部分,得到决策前状态值函数 $V_t^{X^-}(S_t^{X^-}) = \max_{x_t \in X} (F(S_t, x_t) + \xi V_t^X(S_t^X | S_t^{X^-}))$ 和决策后状态值函数 $V_t^X(S_t^X) = E(V_{t+1}^{X^-}(S_{t+1}^{X^-} | S_t^X))$;注意到决策前状态值函数是确定性问题,因此只要解决决策后状态值函数 $V_t^X(S_t^X)$ 的近似,即可通过解决确定性问题来获得最优解。

[0057] 进一步地,步骤S3中,在大量模拟场景下利用基于时序差分TD(1)的策略迭代算法训练值函数,得到收敛的近似值函数,步骤如下:

[0058] S3.1、初始化 $v_t^n(r, W_t)$, $v_t^n(r, W_t)$ 为第n次迭代中t时段第r段的斜率,直接令 $v_t^0(r, W_1) = v_t^0(r, W_2) = \dots = v_t^0(r, W_t) = \dots = v_t^0(r, W_T) = 0$;生成n个训练样本 TS^1, TS^2, \dots, TS^n ;令迭代次数 $g=1$;

[0059] S3.2、从 $t=1, \dots, T$,进行前推计算;

[0060] S3.3、从 $t=T, \dots, 1$,进行回代计算;

[0061] S3.4、令 $g=g+1$,返回步骤S2;当 $g>n$ 时迭代停止,得到收敛的各分段线性凹函数的斜率。

[0062] 进一步地,步骤S3.2的具体步骤如下:

[0063] S3.2.1、令 $t=1$;

[0064] S3.2.2、根据最新的随机变量更新主动配电网系统状态并利用上一次迭代后的各分段斜率 $v_t^{n-1}(r, W_t)$ 求得各决策变量 x_t^n ,决策后的主动配电网系统状态 $S_t^{n,X}$ 包括决策后的可调容量 $R_t^{n,X}$;

[0065] S3.2.3、若 $t>1$,对每个可调容量施加正负扰动,并计算其正负边际效益,第n次迭代中的第i个可调容量 $R_{i,t}^{n,X}$ 的正负边际效益如下:

$$[0066] \begin{cases} \hat{F}_{ec,t}^{n+} = \frac{1}{\delta^+} (F_t(S_t^{n+}, x_t^{n+}) - F_t(S_t, x_t)) \\ \hat{F}_{ec,t}^{n-} = \frac{1}{\delta^-} (F_t(S_t, x_t) - F_t(S_t^{n-}, x_t^{n-})) \end{cases}; \quad (19)$$

[0067] 式中, $\hat{F}_{ec,t}^{n+}$ 、 $\hat{F}_{ec,t}^{n-}$ 分别为第ec个电动汽车子集群可调容量的正负边际效益, δ^+ 、 δ^-

分别为正负扰动量, F_t 为立即成本; S_t^{n+} 、 S_t^{n-} 分别为 t 时段施加正、负扰动后的主动配电网系统状态, x_t^{n+} 、 x_t^{n-} 分别为 t 时段施加正、负扰动后所做的决策;

[0068] S3.2.4、计算每个可调容量施加正负扰动后的资源变化量:

$$[0069] \quad \begin{cases} \Delta R_{ec,t}^{n+,X} = \frac{1}{\delta^+} (R_{ec,t}^{n+,X} - R_{ec,t}^{n,X}) \\ \Delta R_{ec,t}^{n-,X} = \frac{1}{\delta^-} (R_{ec,t}^{n,X} - R_{ec,t}^{n-,X}) \end{cases}; \quad (20)$$

[0070] 其中, $\Delta R_{ec,t}^{n+,X}$ 为施加正扰动后的资源变化量, $\Delta R_{ec,t}^{n-,X}$ 为施加负扰动后的资源变化量; $R_{ec,t}^{n+,X}$ 、 $R_{ec,t}^{n-,X}$ 分别为施加正、负扰动后的资源量;

[0071] S3.2.5、令 $t=t+1$, 若 $t \leq T$, 返回步骤 S2.2; 若 $t > T$, 循环终止。

[0072] 进一步地, 步骤 S3.3 的具体步骤如下:

[0073] S3.3.1、令 $t=T$;

[0074] S3.3.2、计算本次迭代过程中决策前各分段斜率的正负边际值, 正边际值的计算:

$$[0075] \quad \hat{v}_t^{n+}(r, S_t^n) = \begin{cases} \hat{F}_{ec,t}^{n+} + R_{ec,t}^{n+,X} \hat{v}_{t+1}^{n+}(r, S_{t+1}^n), t \neq T \\ \hat{F}_{ec,T}^{n+}, t = T \end{cases}; \quad (21)$$

[0076] 负边际值的计算:

$$[0077] \quad \hat{v}_t^{n-}(r, S_t^n) = \begin{cases} \hat{F}_{ec,t}^{n-} + R_{ec,t}^{n-,X} \hat{v}_{t+1}^{n-}(r, S_{t+1}^n), t \neq T \\ \hat{F}_{ec,T}^{n-}, t = T \end{cases}; \quad (22)$$

[0078] 根据 $\hat{v}_t^{n+}(r, S_t^n)$ 、 $\hat{v}_t^{n-}(r, S_t^n)$ 以及上一次迭代的决策后边际值 $\hat{v}_t^{(n-1)+}(r, S_t^{n-1,X})$ 、 $\hat{v}_t^{(n-1)-}(r, S_t^{n-1,X})$ 求得本次迭代的决策后边际值:

$$[0079] \quad \begin{cases} \hat{v}_{t-1}^{n+}(r, S_{t-1}^{n,X}) = (1-\alpha) \hat{v}_{t-1}^{(n-1)+}(r, S_{t-1}^{n-1,X}) + \alpha \hat{v}_t^{n+}(r, S_t^n) \\ \hat{v}_{t-1}^{n-}(r, S_{t-1}^{n,X}) = (1-\alpha) \hat{v}_{t-1}^{(n-1)-}(r, S_{t-1}^{n-1,X}) + \alpha \hat{v}_t^{n-}(r, S_t^n) \end{cases}; \quad (23)$$

[0080] α 为迭代步长, 取值范围为 $0 \sim 1$;

[0081] S3.3.3、根据 $\hat{v}_{t-1}^{n+}(r, S_{t-1}^{n,X})$ 、 $\hat{v}_{t-1}^{n-}(r, S_{t-1}^{n,X})$, 采用凹自适应值估计法 (The Concave Adaptive Value Estimation Algorithm, CAVE) 对各分段线性凹函数的斜率进行更新;

[0082] S3.3.4、令 $t=t-1$, 若 $t \geq 1$, 返回步骤 S3.2; 若 $t < 1$, 循环终止。

[0083] 进一步地, 步骤 S3.4 中, 利用分段线性函数法将决策后状态值函数用分段线性凹函数表示, 通过不断迭代更新得到收敛的分段线性凹函数的斜率, 从而得到近似的决策后值函数即近似值函数;

[0084] 所述分段线性函数法将决策后状态值函数用分段线性凹函数表示步骤如下:将可调资源用断点 $R = [\rho, 2\rho, \dots, \beta\rho]$ 表示, R 表示断点集合, β 代表段数, ρ 为每段的长度,将分段线性凹函数的斜率向量用有限集合 $v_t(W_t) = [v_t(1, W_t), v_t(2, W_t), \dots, v_t(\beta, W_t)]$ 表示, $v_t(W_t)$ 表示斜率集合, $v_t(\beta, W_t)$ 表示第 β 段的斜率, $[v_t(1, W_t) \geq v_t(2, W_t), \geq \dots \geq v_t(\beta, W_t)]$,保证斜率单调递减;则式(12)可进一步转化为:

$$[0085] \quad x_t = \arg \max_{x_t \in X} (F(S_t, x_t) + \xi \sum_{r=1}^{\beta} v_t(r, W_t) y_{tr}); \quad (17)$$

[0086] 表示当决策 x_t 使得上式取最大值时,即为最优决策;须满足:

$$[0087] \quad \begin{cases} \sum_{r=1}^{\beta} y_{tr} = R_{ec,t}^X \\ 0 \leq y_{tr} \leq \rho, r \in \{1, \dots, \beta\} \end{cases}; \quad (18)$$

[0088] 其中, r 表示第 r 段线性凹函数, ρ 为每段线性凹函数的长度, y_{tr} 为每段线性凹函数的资源量,其中,假设所有的分段线性凹函数的长度都是均匀的。

[0089] 进一步地,步骤S4中,利用训练好的近似值函数对主动配电网实时场景下的随机优化调度问题进行递推求解,得到各时段的近似最优决策,步骤如下:

[0090] S4.1、令 $t=1$;

[0091] S4.2、更新当前时段的外源信息,包括新接入的电动汽车、光伏出力的误差、风机出力的误差以及电价的误差;

[0092] S4.3、利用步骤S3中训练好的近似值函数,根据式(17)计算出 t 时段最优决策;

[0093] S4.4、令 $t=t+1$,若 $t \leq T$,返回步骤S4.2;若 $t > T$,循环终止,得到各时段的近似最优决策。

[0094] 相比于现有技术,本发明具有以下优点:

[0095] 本发明针对主动配电网中多元资源的实时随机优化调度问题,根据强化学习思想,提出了一种能满足主动配电网实时、快速响应的方法,适应多变的配电网调度需求,所提方法在模拟场景中训练过后能快速地收敛并逼近最优解,投入到实时运行中时,一方面能更有效地提升调度收益,另一方面也起到“削峰填谷”的作用。当环境中的随机因素变化更加剧烈时,本发明依旧适用,能接近最优解,且实时调度过程中的计算时间并不受影响,具有较强的鲁棒性。

附图说明

[0096] 图1为本发明实施例中一种基于强化学习的主动配电网实时随机优化调度方法的步骤流程图。

[0097] 图2为本发明实施例中决策前后的电动汽车子集群的状态变化图。

具体实施方式

[0098] 有鉴于此,本发明实施例提供了一种基于强化学习的主动配电网实时随机优化调度方法,满足主动配电网实时、快速响应的方法,适应多变的配电网调度需求。

[0099] 为了使本技术领域的人员更好地理解本发明的技术方案,下面将结合本发明实施

例以及附图,对本发明的技术方案进行清楚、完整地描述,显然,所描述的实施例仅是本发明的一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0100] 实施例:

[0101] 一种基于强化学习的主动配电网实时随机优化调度方法,如图1所示,包括以下步骤:

[0102] S1、建立主动配电网的数学模型,将主动配电网的实时调度描述成一个多阶段随机序贯决策问题;

[0103] 所述主动配电网的数学模型包括潮流约束、储能约束、电动汽车充电约束、分布式电源约束以及主动配电网实时调度的目标函数。

[0104] 所述潮流约束如下:

$$\begin{cases}
 P_{ij}(t) = k_{ij-1} \cdot \frac{(V_i(t) - V_j(t))}{x_{ij}} + k_{ij-2} \cdot \frac{(\delta_i(t) - \delta_j(t))}{x_{ij}} \\
 Q_{ij}(t) = -k_{ij-1} \cdot \frac{(\delta_i(t) - \delta_j(t))}{x_{ij}} + k_{ij-2} \cdot \frac{(V_i(t) - V_j(t))}{x_{ij}} \\
 P_i(t) = \sum_{j=1, j \neq i}^{NB} \frac{k_{ij-2}}{x_{ij}} (\delta_i(t) - \delta_j(t)) + \frac{k_{ij-1}}{x_{ij}} (V_i(t) - V_j(t)) ; \\
 Q_i(t) = \sum_{j=1, j \neq i}^{NB} -\frac{k_{ij-1}}{x_{ij}} (\delta_i(t) - \delta_j(t)) + \frac{k_{ij-2}}{x_{ij}} (V_i(t) - V_j(t)) \\
 k_{ij-1} = \frac{r_{ij} x_{ij}}{r_{ij}^2 + x_{ij}^2}, k_{ij-2} = \frac{x_{ij}^2}{r_{ij}^2 + x_{ij}^2} \\
 -P_{ij, \max} \leq P_{ij} \leq P_{ij, \max}, V_{i, \min} \leq V_i \leq V_{i, \max}
 \end{cases} \quad (1)$$

[0106] 式中, i, j 均为节点编号, P_{ij}, Q_{ij}, r_{ij} 和 x_{ij} 分别为节点 i, j 间的有功传输功率、无功传输功率、电阻和电抗, P_i, Q_i, V_i 和 δ_i 分别为节点 i 的有功功率、无功功率、电压和相角, NB 为支路数; $P_{ij, \max}$ 和 $-P_{ij, \max}$ 分别为节点 i, j 间的最大和最小有功传输功率; $V_{i, \max}$ 和 $V_{i, \min}$ 分别为节点 i 的最大和最小电压;

[0107] 除了与外部电网相连的节点,配电网内其它节点在 t 时段的功率如下:

$$P_{i,t} = P_{i,t}^{DG} + P_{i,t}^{ESS} - P_{i,t}^{CON} - \sum_{n=1}^{N_{EV}} P_{i,n,t}^{EV} ; \quad (2)$$

$$Q_{i,t} = Q_{i,t}^{DG} + Q_{i,t}^{ESS} - Q_{i,t}^{CON} - \sum_{n=1}^{N_{EV}} Q_{i,n,t}^{EV} ; \quad (3)$$

[0110] 式中, $P_{i,t}^{DG}$ 和 $Q_{i,t}^{DG}$ 分别为 t 时段节点 i 上的分布式电源的有功和无功出力;

[0111] $P_{i,t}^{ESS}$ 和 $Q_{i,t}^{ESS}$ 分别为 t 时段节点 i 上的储能的有功和无功出力; $P_{i,t}^{CON}$ 和 $Q_{i,t}^{CON}$ 为 t 时段节点 i 上的常规有功负荷和无功负荷; $P_{i,n,t}^{EV}$ 和 $Q_{i,n,t}^{EV}$ 为 t 时段节点 i 上的电动汽车有功负荷和无功负荷; n 为电动汽车编号, N_{EV} 为电动汽车数目;

[0112] 所述储能约束如下:

$$[0113] \quad E_{a,t+1} = \begin{cases} E_{a,t} - \frac{1}{\eta_a} P_{a,t} & P_{a,t} \geq 0 \\ E_{a,t} - \eta_a P_{a,t} & P_{a,t} < 0 \end{cases}; \quad (4)$$

$$[0114] \quad P_{a,t,\min} \leq P_{a,t} \leq P_{a,t,\max}; \quad (5)$$

$$[0115] \quad E_{a,t,\min} \leq E_{a,t} \leq E_{a,t,\max}; \quad (6)$$

[0116] 式(4)中, a 为储能编号, $E_{a,t}$ 为 t 时段初第 a 个储能的能量, $E_{a,t+1}$ 为 t 时段末、 $t+1$ 时段初第 a 个储能的能量; $P_{a,t}$ 为 t 时段第 a 个储能的充/放电功率,大于0表示放电,小于0表示充电;式(5)和(6)分别为功率和能量的上下限约束, $P_{a,t,\max}$ 和 $P_{a,t,\min}$ 分别为功率上下限, $E_{a,t,\max}$ 和 $E_{a,t,\min}$ 分别为能量上下限;

[0117] 所述分布式电源约束如下:

$$[0118] \quad \begin{cases} P_{d,t+1}^{\text{DG}} - P_{d,t}^{\text{DG}} \leq v_{up,d} \\ P_{d,t}^{\text{DG}} - P_{d,t+1}^{\text{DG}} \leq v_{down,d} \\ P_d^{\text{DG},\min} \leq P_{d,t}^{\text{DG}} \leq P_d^{\text{DG},\max} \end{cases}; \quad (7)$$

[0119] 其中, d 为分布式电源编号, $P_{d,t+1}^{\text{DG}}$ 为第 d 台分布式电源在 $t+1$ 时段中发出的有功功率; $v_{up,d}$ 和 $v_{down,d}$ 分别为第 d 台分布式电源的向上爬坡速率限制和向下爬坡速率限制, $P_d^{\text{DG},\max}$ 和 $P_d^{\text{DG},\min}$ 分别为第 d 台分布式电源输出功率的上限和下限;

[0120] 所述单台电动汽车(EV)的充电约束如下:

$$[0121] \quad \begin{cases} e_{e,t} = e_{e,t_{\text{start}}} + \eta \sum_{k=t_{\text{start}}}^t p_{e,k} \Delta t = e_{e,t-1} + \eta p_{e,t-1} \Delta t \\ e_{e,t_{\text{end}},\min} = e_{e,t_{\text{end}},\max} = e_{\text{exp}} \\ e_{e,t,\min} \leq e_{e,t} \leq e_{e,t,\max} \\ p_{e,t,\max} = \min(p_{\max}, (e_{e,t,\max} - e_{e,t-1}) / \eta / \Delta t) \\ p_{e,t,\min} = \max(0, (e_{e,t,\min} - e_{e,t-1}) / \eta / \Delta t) \\ p_{e,t,\min} \leq p_{e,t} \leq p_{e,t,\max} \end{cases}; \quad (8)$$

[0122] 式中, e 为电动汽车编号, $e_{e,t}$ 为 t 时段初第 e 台电动汽车电池的电量, η 为充电效率, $p_{e,k}$ 表示 k 时段第 e 台电动汽车的充电功率, $e_{e,t,\max}$ 、 $e_{e,t,\min}$ 分别为 t 时刻电动汽车能量的上下边界;能量的上边界表示电动汽车接入电网后便以最大功率充电充至 e_{exp} 这段时间内的电量变化曲线;能量的下边界表示接入后延迟充电直至离开的时刻恰好达到用户期望的电量值的电量变化曲线; e_{exp} 为用户希望离开时电池的电量,一般为电池的最大电量; $p_{e,t,\max}$ 、 $p_{e,t,\min}$ 分别为 t 时段第 e 台电动汽车充电功率的上下限约束, p_{\max} 表示受充电桩和电动汽车本身性质影响的最大充电功率; Δt 为每个时段的时长; t_{end} 表示结束时段;

[0123] 当电动汽车的渗透率过高时,将每台电动汽车作为决策变量进行求解会产生“维数灾”的问题,可利用集群的思想,将一天分为 T 个时段,具有相同离开时段的电动汽车分为同一集群,将集群内单台电动汽车的充电模型叠加,可得到等效的集群模型;

$$[0124] \quad \begin{cases} E_{\min,ec,t} \leq E_{ec,t} \leq E_{\max,ec,t} \\ E_{ec,t} = E_{ec,t_{\text{start}}} + \eta \sum_{k=t_{\text{start}}}^t P_{ec,k} \Delta t = E_{ec,t-1} + \eta P_{ec,t-1} \Delta t \\ P_{ec,t} = \sum_{e=1}^m P_{ec,e,t} \\ P_{\min,ec,t} = \sum_{e=1}^m P_{\min,ec,e,t} \leq P_{ec,t} \leq \sum_{e=1}^m P_{\max,ec,e,t} = P_{\max,ec,t} \\ E_{\min,ec,t} = \sum_{e=1}^m e_{\min,ec,e,t} \\ E_{\max,ec,t} = \sum_{e=1}^m e_{\max,ec,e,t} \end{cases} ; \quad (9)$$

[0125] 式中,ec为电动汽车集群编号, $E_{ec,t}$ 、 $E_{\max,ec,t}$ 、 $E_{\min,ec,t}$ 和 $P_{ec,t}$ 分别为t时段子集群ec的能量、能量上下界和子集群充电总功率; $P_{ec,e,t}$ 为子集群ec内第e辆电动汽车t时段的充电功率;m为t时段子集群ec的电动汽车数量;

[0126] 所述目标函数,为考虑各类资源包括电动汽车、分布式电源的随机性下,通过实时调度决策,使得调度周期内的总收益最大,即将主动配电网的实时调度描述成一个多阶段随机序贯决策问题:

$$[0127] \quad \max_{x_t \in \mathcal{X}_t} \mathbf{E} \left[\sum_{t=1}^T F_t \right]; \quad (10)$$

[0128] 其中, x_t 表示决策变量, \mathcal{X}_t 表示可行域, $\mathbf{E}(\cdot)$ 表示取期望值运算, F_t 表示t时段的调度收益,为调度成本的相反数,具体如下:

$$[0129] \quad F_t = - \left(C_t^{\text{trans}} + \sum_{d \in N_{\text{DG}}} C_{t,d}^{\text{DG}} + \sum_{a \in N_{\text{ESS}}} C_{t,a}^{\text{ESS}} \right); \quad (11)$$

[0130] 其中, C_t^{trans} 、 $C_{t,a}^{\text{ESS}}$ 分别是t时段主动配电网向外部电网购电的成本、t时段第a个储能的运行成本, $C_{t,d}^{\text{DG}}$ 是t时段第d个分布式电源的运行成本,为二次函数; N_{DG} 、 N_{ESS} 分别表示分布式电源、储能的个数。

[0131] S2、提出原问题的动态规划公式,根据主动配电网中电动汽车的特性构造表征当前决策对后续所有时段影响的值函数,并利用决策后状态值函数避免期望值的计算;

[0132] 提出原问题的动态规划公式,即根据贝尔曼的最优性原理,可将多时段的优化决策问题转化成递推问题,将式(10)转化成:

$$[0133] \quad V_t(S_t) = \max_{x_t \in \mathcal{X}} (F(S_t, x_t) + \xi E(V_{t+1}(S_{t+1} | S_t, x_t, W_t))); \quad (12)$$

[0134] 其中, S_t 为主动配电网系统t时段所处的状态, $F(S_t, x_t)$ 为主动配电网系统在 S_t 状态时作了决策 x_t 后产生的实时收益, $V_t(S_t)$ 为主动配电网系统在 S_t 状态的值函数, $V_{t+1}(S_{t+1} | S_t, x_t, W_t)$ 为主动配电网系统在 S_t 状态的前提下,做出决策 x_t 、接收外源信息 W_t 后t+1时刻的值函数,值函数的意义为后续状态对当前状态的影响;

[0135] 所述主动配电网系统t时段所处的状态 S_t 为

$$[0136] \quad S_t = (R_{ec,t}^{\text{EV}}, R_{d,t}^{\text{DG}}, R_{a,t}^{\text{ESS}}, P_{it}, Q_{it}); \quad (13)$$

[0137] 其中, $R_{ec,t}^{\text{EV}}$ 为t时段电动汽车子集群ec的可调资源量, $R_{ec,t}^{\text{EV}} = E_{\max,ec,t} - E_{ec,t}$,即t时段电动汽车子集群ec的可调资源量为t时段电动汽车子集群ec的能量上界与该集群

ec此时的能量之差,电动汽车子集群ec在t时段的能量为电池电量总和; $R_{d,t}^{DG}$ 为t时段分布式电源的可调资源量, $R_{d,t}^{DG} = P_{d,t}^{DG,max} - P_{d,t}^{DG}$,即t时段分布式电源的可调资源量为t时段分布式电源的最大发电功率与实际发电功率之差; $R_{ESSa,t}$ 为储能的可调容量,即剩余电量; $P_{i,t}$ 、 $Q_{i,t}$ 分别为t时段节点i的有功功率、无功功率;

[0138] 决策变量 x_t 为:

$$[0139] \quad x_t = (P_{ec,t}, P_{d,t}^{DG}, P_{a,t}); \quad (14)$$

[0140] 外源信息 W_t 为:

$$[0141] \quad W_t = (\tilde{w}_t^{EV}, \tilde{w}_t^{PV}, \tilde{w}_t^{WT}, \tilde{a}_t); \quad (15)$$

[0142] 外源信息 W_t 表示t时段主动配电网系统更新的各随机变量的预测误差,其中 \tilde{w}_t^{EV} 表示因t时段电动汽车随机接入产生的误差,引起的变化包括各个电动汽车子集群内电动汽车的数量、集群能量轨迹上下边界、集群的充电功率极限; \tilde{w}_t^{PV} 、 \tilde{w}_t^{WT} 和 \tilde{a}_t 分别表示t时段光伏出力的误差、风机出力的误差和电价的误差。

[0143] 根据主动配电网中电动汽车的特性构造表征当前决策对后续所有时段影响的值函数,并利用决策后状态值函数避免期望值的计算;

[0144] 首先定义状态转移方程:

$$[0145] \quad S_t = f(S_{t-1}, x_{t-1}, W_t); \quad (16)$$

[0146] 表示t时段的主动配电网系统状态 S_t 由t-1的状态 S_{t-1} 、t-1时段所做的决策 x_{t-1} 、t时段的外源信息 W_t 共同决定;

[0147] 接着将主动配电网系统状态 S_t 分为决策前状态 $S_t^{X^-}$ 和决策后状态 S_t^X 两个阶段;决策前,主动配电网系统接收外源信息 W_t ,改变状态, $S_t^{X^-} = S_{t-1}^{X^-} + f^{X^-}(W_t)$;做出决策后、未接受外源信息前,主动配电网系统的状态进一步变化 $S_t^X = S_t^{X^-} + f^X(x_t)$;

[0148] 根据电动汽车特性进行设计,如图2所示,决策前每个电动汽车子集群的电池电量 $E_{ec,t}$ (实线)不变,但观测到新的电动汽车的随机接入,即 \tilde{w}_t^{EV} ,导致电动汽车集群的能量上下边界 $E_{ec,t,max}$ 、 $E_{ec,t,min}$ 发生变化,从而电动汽车子集群的可调资源量由 $R_{ec,t-1}^X$ 变化至 $R_{ec,t}^X$;决策后,每个电动汽车集群的电池电量 $E_{ec,t}$ 根据所作的决策发生改变,可调资源量变成 $R_{Xec,t}$,实现主动配电网系统状态的转移;

[0149] 根据前述内容将主动配电网系统状态分成决策前与决策后状态,则式(12)可分成两部分,得到决策前状态值函数 $V_t^{X^-}(S_t^{X^-}) = \max_{x_t \in X} (F(S_t, x_t) + \xi V_t^X(S_t^X | S_t^{X^-}))$ 和决策后状态值函数 $V_t^X(S_t^X) = E(V_{t+1}^{X^-}(S_{t+1}^{X^-} | S_t^X))$;注意到决策前状态值函数是确定性问题,因此只要解决决策后状态值函数 $V_t^X(S_t^X)$ 的近似,即可通过解决确定性问题来获得最优解。

[0150] S3、在大量模拟场景下利用基于时序差分TD(1)的策略迭代算法训练值函数,得到近似值函数;,步骤如下:

[0151] S3.1、初始化 $v_t^n(r, W_t)$, $v_t^n(r, W_t)$ 为第n次迭代中t时段第r段的斜率, 直接令 $v_1^0(r, W_1) = v_2^0(r, W_2) = \dots = v_t^0(r, W_t) = \dots = v_T^0(r, W_T) = 0$; 生成n个训练样本 TS^1, TS^2, \dots, TS^n ; 令迭代次数 $g=1$;

[0152] S3.2、从 $t=1, \dots, T$, 进行前推计算; 具体步骤如下:

[0153] S3.2.1、令 $t=1$;

[0154] S3.2.2、根据最新的随机变量更新主动配电网系统状态并利用上一次迭代后的各分段斜率 $v_t^{n-1}(r, W_t)$ 求得各决策变量 x_t^n , 决策后的主动配电网系统状态 $S_t^{n,X}$ 包括决策后的可调容量 $R_t^{n,X}$;

[0155] S3.2.3、若 $t>1$, 对每个可调容量施加正负扰动, 并计算其正负边际效益, 第n次迭代中的第i个可调容量 $R_{i,t}^{n,X}$ 的正负边际效益如下:

$$[0156] \quad \begin{cases} \hat{F}_{ec,t}^{n+} = \frac{1}{\delta^+} (F_t(S_t^{n+}, x_t^{n+}) - F_t(S_t, x_t)) \\ \hat{F}_{ec,t}^{n-} = \frac{1}{\delta^-} (F_t(S_t, x_t) - F_t(S_t^{n-}, x_t^{n-})) \end{cases}; \quad (19)$$

[0157] 式中, $\hat{F}_{ec,t}^{n+}$ 、 $\hat{F}_{ec,t}^{n-}$ 分别为第ec个电动汽车子集群可调容量的正负边际效益, δ^+ 、 δ^- 分别为正负扰动量, F_t 为立即成本; S_t^{n+} 、 S_t^{n-} 分别为t时段施加正、负扰动后的主动配电网系统状态, x_t^{n+} 、 x_t^{n-} 分别为t时段施加正、负扰动后所做的决策;

[0158] S3.2.4、计算每个可调容量施加正负扰动后的资源变化量:

$$[0159] \quad \begin{cases} \Delta R_{ec,t}^{n+,X} = \frac{1}{\delta^+} (R_{ec,t}^{n+,X} - R_{ec,t}^{n,X}) \\ \Delta R_{ec,t}^{n-,X} = \frac{1}{\delta^-} (R_{ec,t}^{n,X} - R_{ec,t}^{n-,X}) \end{cases}; \quad (20)$$

[0160] 其中, $\Delta R_{ec,t}^{n+,X}$ 为施加正扰动后的资源变化量, $\Delta R_{ec,t}^{n-,X}$ 为施加负扰动后的资源变化量; $R_{ec,t}^{n+,X}$ 、 $R_{ec,t}^{n-,X}$ 分别为施加正、负扰动后的资源量;

[0161] S3.2.5、令 $t=t+1$, 若 $t \leq T$, 返回步骤S2.2; 若 $t > T$, 循环终止。

[0162] S3.3、从 $t=T, \dots, 1$, 进行回代计算; 具体步骤如下:

[0163] S3.3.1、令 $t=T$;

[0164] S3.3.2、计算本次迭代过程中决策前各分段斜率的正负边际值, 正边际值的计算:

$$[0165] \quad v_t^{n+}(r, S_t^n) = \begin{cases} \hat{F}_{ec,t}^{n+} + R_{ec,t}^{n+,X} v_{t+1}^{n+}(r, S_{t+1}^n), & t \neq T \\ \hat{F}_{ec,T}^{n+}, & t = T \end{cases}; \quad (21)$$

[0166] 负边际值的计算:

$$[0167] \quad \hat{v}_t^{n-} (r, S_t^n) = \begin{cases} \hat{F}_{ec,t}^{n-} + R_{ec,t}^{n-,X} \hat{v}_{t+1}^{n-} (r, S_{t+1}^n), & t \neq T \\ \hat{F}_{ec,T}^{n-}, & t = T \end{cases}; \quad (22)$$

[0168] 根据 $\hat{v}_t^{n+}(r, S_t^n)$ 、 $\hat{v}_t^{n-}(r, S_t^n)$ 以及上一次迭代的决策后边际值 $\hat{v}_t^{(n-1)+}(r, S_t^{n-1,X})$ 、 $\hat{v}_t^{(n-1)-}(r, S_t^{n-1,X})$ 求得本次迭代的决策后边际值:

$$[0169] \quad \begin{cases} \hat{v}_{t-1}^{n+}(r, S_{t-1}^{n,X}) = (1-\alpha) \hat{v}_{t-1}^{(n-1)+}(r, S_{t-1}^{n-1,X}) + \alpha \hat{v}_t^{n+}(r, S_t^n) \\ \hat{v}_{t-1}^{n-}(r, S_{t-1}^{n,X}) = (1-\alpha) \hat{v}_{t-1}^{(n-1)-}(r, S_{t-1}^{n-1,X}) + \alpha \hat{v}_t^{n-}(r, S_t^n) \end{cases}; \quad (23)$$

[0170] α 为迭代步长, 取值范围为 $0 \sim 1$;

[0171] S3.3.3、根据 $\hat{v}_{t-1}^{n+}(r, S_{t-1}^{n,X})$ 、 $\hat{v}_{t-1}^{n-}(r, S_{t-1}^{n,X})$, 采用凹自适应值估计法 (The Concave Adaptive Value Estimation Algorithm, CAVE) 对各分段线性凹函数的斜率进行更新;

[0172] S3.3.4、令 $t = t - 1$, 若 $t \geq 1$, 返回步骤 S3.2; 若 $t < 1$, 循环终止。

[0173] S3.4、令 $g = g + 1$, 返回步骤 S2; 当 $g > n$ 时迭代停止, 得到收敛的各分段线性凹函数的斜率;

[0174] 其中, 利用分段线性函数法将决策后状态值函数用分段线性凹函数表示, 通过不断迭代更新得到收敛的分段线性凹函数的斜率, 从而得到近似的决策后值函数即近似值函数;

[0175] 所述分段线性函数法将决策后状态值函数用分段线性凹函数表示步骤如下: 将可调资源用断点 $R = [\rho, 2\rho, \dots, \beta\rho]$ 表示, R 表示断点集合, β 代表段数, ρ 为每段的长度, 将分段线性凹函数的斜率向量用有限集合 $v_t(W_t) = [v_t(1, W_t), v_t(2, W_t), \dots, v_t(\beta, W_t)]$ 表示, $v_t(W_t)$ 表示斜率集合, $v_t(\beta, W_t)$ 表示第 β 段的斜率, $[v_t(1, W_t) \geq v_t(2, W_t), \geq \dots \geq v_t(\beta, W_t)]$, 保证斜率单调递减; 则式 (12) 可进一步转化为:

$$[0176] \quad x_t = \arg \max_{x_t \in X} (F(S_t, x_t) + \xi \sum_{r=1}^{\beta} v_t(r, W_t) y_{tr}); \quad (17)$$

[0177] 表示当决策 x_t 使得上式取最大值时, 即为最优决策; 须满足:

$$[0178] \quad \begin{cases} \sum_{r=1}^{\beta} y_{tr} = R_{ec,t}^X \\ 0 \leq y_{tr} \leq \rho, r \in \{1, \dots, \beta\} \end{cases}; \quad (18)$$

[0179] 其中, r 表示第 r 段线性凹函数, ρ 为每段线性凹函数的长度, y_{tr} 为每段线性凹函数的资源量, 其中, 假设所有的分段线性凹函数的长度都是均匀的。

[0180] S4、利用训练好的近似值函数对主动配电网实时场景下的随机优化调度问题进行递推求解, 得到各时段的近似最优决策; 步骤如下:

[0181] S4.1、令 $t = 1$;

[0182] S4.2、更新当前时段的外源信息,包括新接入的电动汽车、光伏出力的误差、风机出力的误差以及电价的误差;

[0183] S4.3、利用步骤S3中训练好的近似值函数,根据式(17)计算出t时段最优决策;

[0184] S4.4、令 $t=t+1$,若 $t \leq T$,返回步骤S4.2;若 $t > T$,循环终止,得到各时段的近似最优决策。

[0185] 本发明中,针对主动配电网中多元资源的实时随机优化调度问题,根据强化学习思想,提出了一种能满足主动配电网实时、快速响应的方法,适应多变的配电网调度需求,所述方法在模拟场景中训练过后能快速地收敛并逼近最优解,投入到实时运行中时,一方面能更有效地提升调度收益,另一方面也起到“削峰填谷”的作用。当环境中的随机因素变化更加剧烈时,本发明依旧适用,能接近最优解,且实时调度过程中的计算时间并不受影响,具有较强的鲁棒性。

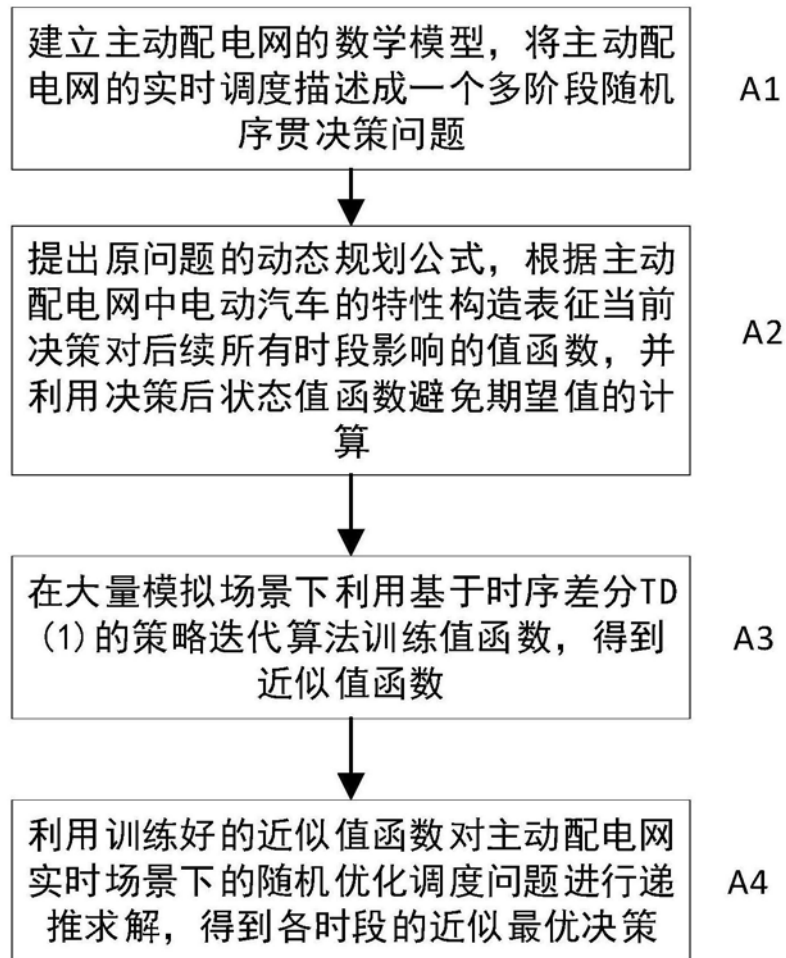


图1

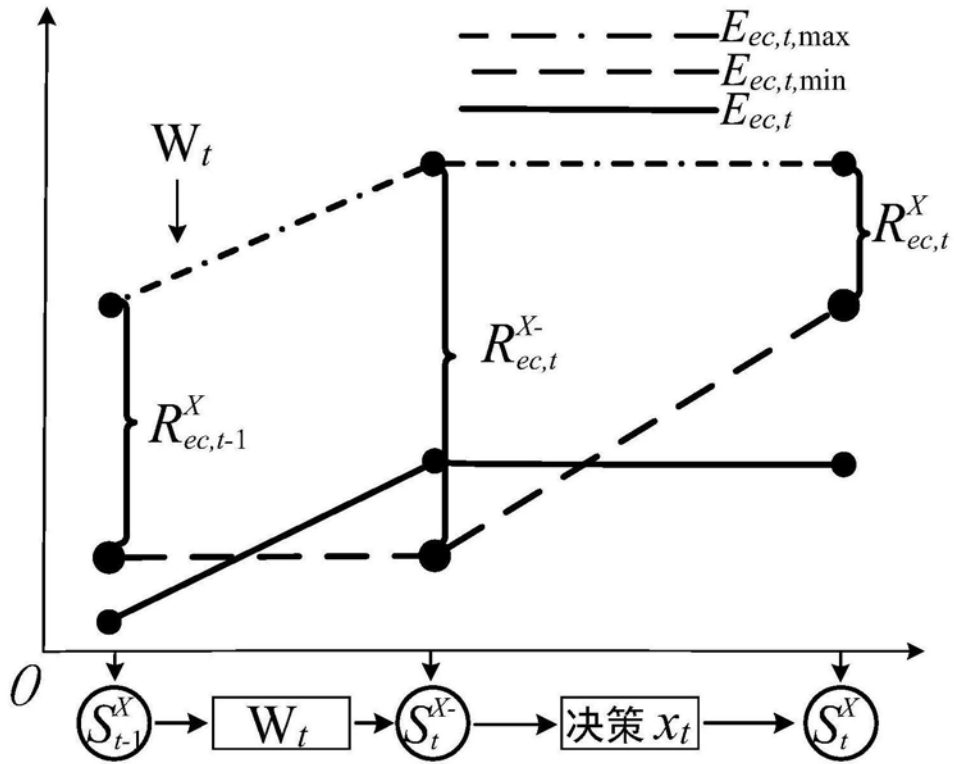


图2